Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application. No changes to the claims are made herein.

Listing of Claims:

1. **(Previously Presented)** A method of providing pattern recognition, said method comprising the steps of:

inputting a pattern;

transforming the input pattern to provide a set of at least one feature for a classifier;

said transforming step comprising the step of minimizing the probability of subsequent misclassification of the at least one feature in the classifier;

said minimizing step comprising:

developing an objective function; and

optimizing the objective function through gradient descent,

wherein said minimizing step is performed non-incrementally.

2. **(Original)** The method according to Claim 1, wherein said minimizing step comprises maximizing an average pairwise divergence.

3. **(Cancelled)**

4. **(Original)** The method according to Claim 1, further comprising the step of querying whether the optimized objective function converges.

5. **(Original)** The method according to Claim 4, further comprising the step of repeating said optimizing step if the optimized objective function does not converge.

6. **(Original)** The method according to Claim 1, wherein said pattern recognition is speech recognition.

7. **(Previously Presented)** Apparatus for providing pattern recognition, said apparatus comprising:

an input interface for inputting a pattern;

a transformer for transforming the input pattern to provide a set of at least one feature for a classifier;

said transformer being adapted to minimize the probability of subsequent misclassification of the at least one feature in the classifier;

said transformer further being adapted to:

develop an objective function; and

optimize the objective function through gradient descent,

wherein said minimization is performed non-incrementally.

8. **(Original)** The apparatus according to Claim 7, wherein said transformer is

adapted to minimize the probability of subsequent misclassification of the at least one

feature in the classifier via maximizing an average pairwise divergence.

9. **(Cancelled)**

10. **(Original)** The apparatus according to Claim 7, wherein said transformer is

further adapted to query whether the optimized objective function converges.

11. **(Original)** The apparatus according to Claim 10, wherein said transformer is

further adapted to repeat optimization of the objective function if the optimized objective

function does not converge.

12. **(Original)** The apparatus according to Claim 7, wherein said pattern

recognition is speech recognition.

13. **(Previously Presented)** A program storage device readable by machine,

tangibly embodying a program of instructions executable by the machine to perform

method steps for providing pattern recognition, said method comprising the steps of:

inputting a pattern;

transforming the input pattern to provide a set of at least one feature for a

classifier;

said transforming step comprising the step of minimizing the probability of

subsequent misclassification of the at least one feature in the classifier;

said minimizing step comprising:

developing an objective function; and

optimizing the objective function through gradient descent,

wherein said minimization is performed non-incrementally.

14. **(Previously Presented)** The method according to claim 1, wherein said objective function is an average pairwise divergence related to the probability of misclassification of a projected space based on classes having uniform prior probabilities.

15. **(Previously Presented)** The method according to claim 14, wherein said gradient descent comprises matrix differentiation in which a multidimensional matrix is analyzed non-incrementally.

16. **(Previously Presented)** The method according to claim 14, further comprising the steps of maximizing said average pairwise divergence via numerical optimization routines.

17. **(Previously Presented)** The method according to claim 1, wherein said objective function comprises means, covariances, and prior probabilities.

18. **(Previously Presented)** The method according to claim 1, wherein said objective function is expressed by the following equation:

$$D_\theta = \frac{1}{C(C-1)} trace \left\{ \sum_{i=1}^{C} \left(\theta \Sigma_i \theta^T \right)^{-1} \theta S_i \theta^T \right\} - p$$

where $S_i = \sum_{j \neq i} \Sigma_j + (\mu_i - \mu_j)(\mu_i - \mu_j)^T$, $i = 1, \ldots, C$.